

개인의 특성 기반 패션 추천 시스템



저자 1 (201924532 이성훈)

저자 2 (201924451 김태훈)

조준수

목 차

| | |
|---|---|
| 1. 서론 | 1 |
| 1.1 연구 배경 | 1 |
| 1.2 기존 문제점 | 1 |
| 1.3 연구 목표 | 1 |
| 2. 연구 배경 | 2 |
| 2.1 퍼스널 컬러와 얼굴형, 체형의 중요성 | 2 |
| 2.2 개인화 서비스의 수요 증가와 패션 추천 시스템의 한계 | 3 |
| 3. 연구 내용 | 4 |
| 3.1 시스템 설계 | 4 |
| 3.1.1 사용자 입력 및 데이터 처리 | 4 |
| 3.1.2 데이터 전처리 및 분석 방법 | 4 |
| 3.1.3 전체 구상도 | 4 |
| 3.2 퍼스널 컬러 측정 | 5 |
| 3.2.1 데이터 전처리 | 5 |
| 3.2.2 피부 영역 RGB 값 추출 | 5 |
| 3.2.3 색상 공간 변환 및 퍼스널 컬러 분류 | 5 |
| 3.3 얼굴형 측정 | 6 |
| 3.3.1 데이터 전처리 | 6 |
| 3.3.2 모델 구성 | 6 |
| 3.4 체형 측정 | 6 |
| 3.4.1 KeyPoint 추출 | 6 |
| 3.4.2 허리 위치 추측 | 7 |

| | |
|---|----|
| 3.4.3 가슴 위치 추측 | 8 |
| 3.4.4 배경 제거 및 길이 측정 | 8 |
| 3.5 RandomForest를 활용한 추천시스템 | 9 |
| 3.5.1 데이터 전처리 | 9 |
| 3.5.2 Random Forest를 활용한 학습 | 10 |
| 3.5.3 추천 품질에 대한 피드백 | 11 |
| 3.6 콘텐츠 기반 협업필터링 추천시스템 | 12 |
| 3.6.1 데이터 전처리 | 12 |
| 3.6.2 Tf-Idf Vectorizing | 12 |
| 3.6.3 코사인 유사도 계산 및 옷 특성 연결 | 12 |
| 4. 연구 결과 분석 및 평가 | 13 |
| 4.1 퍼스널 컬러 측정 모델 | 13 |
| 4.2 얼굴형 측정 모델 | 14 |
| 4.2.1 EfficientNetB4를 제외한 나머지만 학습 | 14 |
| 4.2.2 전체 학습 | 14 |
| 4.3 체형 측정 모델 | 15 |
| 4.4 추천시스템 | 15 |
| 5. 결론 및 향후 연구 방향 | 17 |
| 6. 참고 문헌 | 17 |

1. 서론

1.1 연구 배경

현대 사회에서 패션은 개인의 개성을 표현하는 중요한 수단이다. 그러나 모든 사람에게 적합한 패션을 찾기란 쉽지 않다. 특히 퍼스널 컬러, 얼굴형, 체형 등 신체적 특성에 기반한 맞춤형 패션 추천 시스템은 아직 일반적이지 않으며, 이로 인해 사용자들은 자신에게 적합한 패션을 쉽게 찾지 못하는 어려움이 있다. 최근 개인 맞춤형 서비스에 대한 수요가 증가하고 있으며, 이를 바탕으로 신체적 특성을 고려한 패션 추천 시스템을 개발하는 것이 중요해지고 있다.

1.2 기존 문제점

기존의 패션 추천 시스템들은 주로 트렌드나 인기 아이템을 기반으로 추천하는 경우가 많으며, 개개인의 신체적 특성을 충분히 고려하지 않는 경우가 많다. 이는 사용자에게 최적의 추천을 제공하지 못하며, 사용자는 자신에게 맞지 않는 옷을 추천받을 가능성이 크다.

1.3 연구 목표

본 연구의 목표는 사용자의 퍼스널 컬러, 얼굴형, 체형 등 신체적 특성을 기반으로 한 패션 추천 시스템을 개발하는 것이다. 이를 통해 사용자에게 맞춤형 패션을 추천함으로써 패션 선택의 어려움을 줄이고, 사용자 만족도를 높이는 것이 목표이다.

2. 연구 배경

2.1 퍼스널 컬러와 얼굴형, 체형의 중요성

퍼스널 컬러는 개개인이 타고난 신체적 색채(피부, 머리카락, 눈동자 등)를 의미하며, 개인의 전체적인 외모 이미지를 좌우하는 중요한 요소이다. 캐롤 잭슨이 제안한 퍼스널 컬러 색채 시스템에 따르면 사람의 색상은 크게 봄, 여름, 가을, 겨울의 4가지 유형으로 구분된다. 퍼스널 컬러에 맞는 옷을 착용할 경우, 개인의 피부가 더 생기 있어 보이며 전체적인 인상이 밝고 매력적으로 표현된다. 반면, 잘 맞지 않는 색상을 입으면 얼굴이 창백해 보이거나 인상이 둔해질 수 있어, 퍼스널 컬러에 맞춘 패션 선택이 중요하다 [1].

얼굴형은 사람마다 고유한 뼈 구조와 피부 조직, 얼굴 근육 등에 따라 달라진다. 흔히 하트형, 긴형, 타원형, 둥근형, 사각형 등으로 분류되며, 얼굴형에 맞는 의상, 액세서리, 헤어스타일 등을 선택하는 것은 얼굴의 균형을 맞추고 매력을 극대화하는 데 매우 중요하다. 예를 들어, 긴 얼굴형의 경우 넓은 라펠이나 오버사이즈 옷이 더 어울리며, 둥근 얼굴형은 길게 떨어지는 스타일이나 브이넥 디자인이 인상을 개선하는 데 효과적이다 [2]. 따라서, 얼굴형을 정확히 분석하여 맞춤형 패션을 추천하는 시스템은 매우 유용할 수 있다.

체형은 신체의 전체적인 구조와 비율을 말하며, 각 개인의 뼈대와 근육 분포, 지방의 비율에 따라 크게 달라진다. 남성은 주로 삼각형, 역삼각형, 사각형, 타원형, 사다리꼴 등의 체형으로 분류되고, 여성은 모래시계형, 역삼각형, 직사각형, 삼각형, 둥근형 등으로 분류된다. 체형에 맞는 옷을 선택하면 체형 보완이 가능하며, 스타일링에 대한 자신감도 상승할 수 있다. 예를 들어, 역삼각형 체형의 경우, 허리를 강조하는 디자인이 어울리고, 둥근형 체형은 플레어 스커트나 브이넥 상의가 체형을 보완하는 데 도움을 준다 [3]. 체형에 맞는 의상을 추천해 주는 시스템은 고객의 만족도를 크게 높일 수 있다.

이처럼 퍼스널 컬러, 얼굴형, 체형에 맞는 패션을 추천하는 시스템은 단순히 외적인 아름다움을 넘어서, 개인의 자존감 향상 및 자신감을 증대시킬 수 있는 강력한 도구가 될 수 있다. 특히 이러한 요소를 제대로 반영한 시스템은 사용자에게 맞춤형 패션 스타일을 제안함으로써 높은 만족도를 제공할 수 있다.

2.2 개인화 서비스의 수요 증가와 패션 추천 시스템의 한계

최근 몇 년간 개인 맞춤형 서비스에 대한 수요가 전 세계적으로 급증하고 있다. 소비자들은 이제 단순히 제품을 구매하는 것을 넘어, 자신에게 딱 맞는 제품과 경험을 원한다. 이는 패션 산업에도 뚜렷하게 반영되고 있으며, 많은 패션 브랜드들이 개인화된 쇼핑 경험을 제공하는 데 주력하고 있다.

2023년 소비 트렌드 보고서에 따르면, 소비자들은 개성 표현과 자신에게 딱 맞는 상품을 찾는 데 큰 관심을 가지고 있으며, 특히 20~30대 소비자 사이에서 맞춤형 서비스에 대한 수요가 높아지고 있다 [4].

그러나 현재의 패션 추천 시스템은 주로 대중적인 트렌드나 사용자의 구매 이력을 바탕으로 작동하는 경우가 많다. 이러한 시스템은 개인의 신체적 특성(퍼스널 컬러, 얼굴형, 체형 등)을 충분히 고려하지 않기 때문에, 사용자에게 적합한 패션을 추천하는 데 한계가 있다. 또한, 현재 시스템들은 사용자 피드백을 제대로 반영하지 못하는 경우가 많다. 사용자로부터 별점이나 리뷰를 받아도 이를 효과적으로 모델에 반영하지 않아 추천 품질을 향상시키지 못하는 문제도 존재한다. 지속적인 피드백 반응을 통해 추천 알고리즘의 정확도를 높이는 것은 필수적이며, 이는 사용자의 신뢰도와 만족도를 동시에 높일 수 있는 중요한 요소다.

이와 같은 한계점을 보완하기 위해, 본 연구는 신체적 특성을 반영한 맞춤형 패션 추천 시스템을 개발하여, 개인화된 패션 경험을 제공하고자 한다. 이 시스템은 사용자가 사진을 업로드하거나 신체적 정보를 입력하면, 이를 자동으로 분석하여 개개인의 특성에 맞는 패션을 추천함으로써 기존의 한계를 극복할 수 있을 것이다.

3. 연구 내용

3.1 시스템 설계

3.1.1 사용자 입력 및 데이터 처리

사용자는 두 가지 방법으로 신체적 특성을 제공할 수 있다. 첫째, 자신의 얼굴 또는 전신 사진을 업로드하는 방식이며, 둘째, 직접 신체적 정보를 입력하는 방식이다. 사진 업로드 방식을 사용하면 딥러닝 기반 이미지 분석 모델을 통해 사용자의 신체적 특성(퍼스널 컬러, 얼굴형, 체형)을 자동으로 추출한다.

이 과정에서 사용자의 사진은 HTTP 프로토콜을 통해 백엔드 서버로 전송된다. 데이터는 JSON 형식 또는 이미지 파일 형식으로 전송되며, 사용자가 업로드한 사진은 별도로 저장하지 않는다. 이를 통해 사용자의 개인정보 보호를 강화하였다.

3.1.2 데이터 전처리 및 분석 방법

사용자가 업로드한 사진은 먼저 이미지 전처리 과정을 거친다. 전처리 단계에서는 얼굴이나 신체의 불필요한 부분을 제거하고, 필요한 부분만을 추출하여 분석한다. 이를 위해 OpenCV와 dlib, MediaPipe를 사용하여 얼굴과 체형의 주요 부분을 정확하게 탐지하고, 이를 바탕으로 피부 영역, 얼굴형, 체형 정보를 추출한다.

3.1.3 전체 구상도

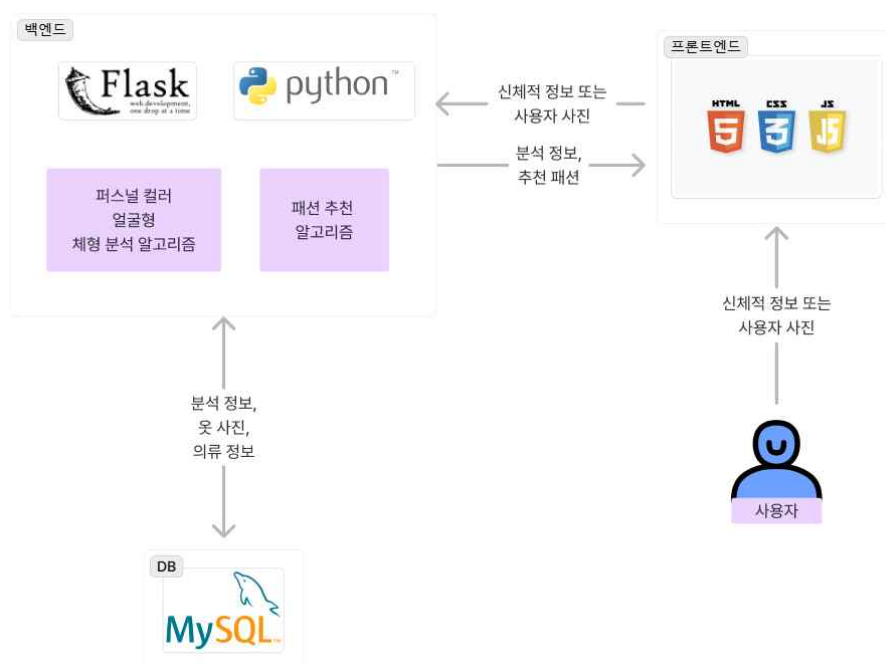


그림1. 전체 구상도

3.2 퍼스널 컬러 측정

3.2.1 데이터 전처리

사용된 데이터셋은 AIHub의 한국인 안면 이미지로, 약 3만장의 이미지로 구성된다. 이 중에서 정면 사진만을 사용하여, 약 5000개의 이미지를 사용하였다. 각 이미지의 얼굴의 피부 영역만을 추출하여 정확한 퍼스널 컬러 측정을 진행하였다.

얼굴에서 피부 영역만을 추출하기 위해 딥러닝 기반의 얼굴 탐지 및 세분화 모델을 적용하였다. facer 라이브러리의 retinaface/mobilenet 모델을 사용하여 얼굴을 감지한 후, farf/lapa/448 모델을 활용하여 얼굴을 세분화했다. 이 과정에서 피부 영역을 정확히 구분하기 위해 얼굴의 각 부분(눈, 코, 입, 이마, 피부 등)을 확률 맵으로 분할하고, 피부 부분만을 선택하였다.

3.2.2 피부 영역 RGB 값 추출

세분화된 확률 맵에서 피부 부분에 해당하는 이진 마스크를 생성하였다. 이 마스크는 피부 영역의 픽셀만 선택할 수 있도록 해주며, 이를 통해 이미지에서 피부의 RGB 값을 추출하였다. 추출된 RGB 값은 노이즈를 제거하고 데이터의 왜곡을 최소화하기 위해 1사분위(Q1)와 3사분위(Q3) 사이의 값들만 사용하여 필터링된 평균값을 도출하였다.

3.2.3 색상 공간 변환 및 퍼스널 컬러 분류

추출된 평균 RGB 값은 HSV 및 Lab 색상 공간으로 변환하였다. 특히, HSV 색상 공간에서의 V(Value)와 S(Saturation) 값, 그리고 Lab 색상 공간에서의 b 채널 값을 활용하여 분석하였다.

이후, K-means 클러스터링 알고리즘을 사용하여 추출된 피부색에 따라 사용자의 퍼스널 컬러를 예측하는 모델을 구축하였다. 이 모델은 이소영(2019)의 논문을 참고하여 설계되었으며, 해당 논문에서 제시한 계절별 대표 피부색의 명도, 채도, 황색도를 기준으로 봄, 여름, 가을, 겨울로 나누는 방식을 사용하였다 [5].

| 피부색 | | | |
|-------------------|-------------------|-------------------|-------------------|
| 봄 | 여름 | 가을 | 겨울 |
| 고명도, 고채도, 고황색도 | 고명도, 저채도, 저황색도 | 저명도, 저채도, 고황색도 | 저명도, 고채도, 저황색도 |

표 1. 퍼스널컬러 유형 별 피부색 특징

3.3 얼굴형 측정

3.3.1 데이터 전처리

사용된 데이터셋은 Kaggle의 Face Shape Dataset으로, 총 5000개의 이미지로 구성된다. 각 얼굴형(하트형, 긴형, 타원형, 둥근형, 사각형)별로 800장의 훈련용 이미지와 200장의 테스트용 이미지가 포함되어 있다. 해당 이미지들은 OpenCV에서 제공하는 CascadeClassifier를 사용하여 얼굴부분만을 추출하였고, 이미지 크기를 380x380 크기로 조정하였다. 이는 EfficientNetB4가 380x380의 input shape를 가지고 있기 때문이다. 5000장의 이미지는 훈련하기에 부족하다고 판단되어 데이터를 증강하였으며, 원본 사진을 수평으로 뒤집은 사진과 $(-20, 20)$ 도 범위에서 회전시킨 이미지들을 추가하였다.

또한, 사진의 픽셀값 범위는 0~255 범위에 있으므로, 전체 이미지를 R, G, B 평균값과 표준편차를 기반으로 아래와 같이 정규화하였다.

$$p_i = \frac{p_i - mean_i}{std_i}, i \in r, g, b$$

여기서 $p_i, mean_i, std_i$ 는 각각 i 색상 픽셀 중 하나, 전체 사진의 i 색상 값 전체 평균, 전체 사진의 i 색상 값 전체 표준편차이다.

3.3.2 모델 구성

모델 구성은 ImageNet 데이터로 사전학습된 EfficientNetB4 네트워크를 기반으로 구성하였다. EfficientNetB4에서 GlobalAveragePooling2D를 적용하고 출력되도록 설정하였으며, EfficientNetB4에서 나온 출력값을 Dense층을 통해 5개의 얼굴형으로 분류하였다.

3.4 체형 측정

3.4.1 KeyPoint 추출

체형 측정은 MediaPipe 라이브러리의 Pose Landmark 기능을 사용하여 신체 부위의 KeyPoint 좌표를 추출하는 방식으로 이루어진다. 이 과정에서 어깨, 엉덩이, 무릎 등의 좌표를 추출하고, 각 부위별로 비율을 계산하여 허리의 위치를 예측하였다. 한국인의 신체 데이터를 활용하여 각 신체치수 대비 허리높이 비율을 계산하고, 이를 기반으로 허리 위치를 추정하였다.

3.4.2 허리 위치 추측

한국인의 신체 데이터를 정리한 사이즈 코리아의 데이터를 사용하였다. 해당 데이터는 한국인의 연령별 및 성별 어깨높이, 엉덩이높이, 허리높이, 무릎높이 등이 정리되어있다. 체형별로 정리되어있어 각 체형의 평균 어깨높이, 엉덩이높이, 허리높이, 무릎높이를 구하고, 어깨높이 대비 허리높이, 엉덩이높이 대비 허리높이, 무릎높이 대비 허리높이, (어깨높이+엉덩이높이) 대비 허리높이, (어깨높이+엉덩이높이+무릎높이) 대비 허리높이를 구하여 어느 것이 허리높이를 구하는데 적합한지 알아보았다.

| 단위: mm, 평균 (표준편차) | 20대 남성 | 30대 남성 | 40대 남성 | 50대 남성 | 60대 남성 |
|----------------------|---------------------|--------------------|--------------------|--------------------|--------------------|
| 평균 어깨높이 | 1398.67 | 1385.75 | 1364 | 1348.75 | 1339 |
| 평균 엉덩이높이 | 849.33 | 834.5 | 819.25 | 805 | 804 |
| 평균 허리높이 | 1070.33 | 1054.24 | 1033 | 1024.25 | 1012.75 |
| 평균 무릎높이 | 443 | 438.25 | 431.5 | 422.25 | 424.5 |
| 허리높이/어깨높이 | 0.765 (0.000042) | 0.761 (0.00305) | 0.757 (0.00648) | 0.759 (0.01437) | 0.756 (0.01389) |
| 허리높이/엉덩이높이 | 1.26 (0.00393) | 1.263 (0.0096) | 1.261 (0.01699) | 1.273 (0.04887) | 1.261 (0.04218) |
| 허리높이/무릎높이 | 2.416 (0.01572) | 2.406 (0.02015) | 2.395 (0.0525) | 2.427 (0.08395) | 2.388 (0.09247) |
| 허리높이/어깨+ 엉덩이) | 0.476 (0.00072) | 0.475 (0.0023) | 0.473 (0.00475) | 0.476 (0.01182) | 0.473 (0.01073) |
| 허리높이/(어깨+ 엉덩이+무릎) | 0.398 (0.00091) | 0.397 (0.00202) | 0.395 (0.00471) | 0.398 (0.01047) | 0.395 (0.00995) |

표 2. 남성의 연령별 신체치수와 각 신체치수 대비 허리높이

| 단위: mm, 평균 (표준편차) | 20대 여성 | 30대 여성 | 40대 여성 | 50대 여성 | 60대 여성 |
|----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| 평균 어깨높이 | 1289.5 | 1268.25 | 1260 | 1249.25 | 1223 |
| 평균 엉덩이높이 | 778.75 | 756.75 | 748.25 | 744.56 | 729.5 |
| 평균 허리높이 | 995 | 971.75 | 963.25 | 951.31 | 927.5 |
| 평균 무릎높이 | 407.75 | 398 | 397.25 | 393.81 | 386.5 |
| 허리높이/어깨높이 | 0.772 (0.00317) | 0.766 (0.00555) | 0.764 (0.00509) | 0.76 (0.00733) | 0.758 (0.00971) |
| 허리높이/엉덩이높이 | 1.278 (0.00395) | 1.284 (0.00729) | 1.287 (0.00729) | 1.274 (0.01698) | 1.271 (0.01685) |
| 허리높이/무릎높이 | 2.44 (0.00729) | 2.442 (0.0117) | 2.425 (0.00835) | 2.413 (0.00431) | 2.4 (0.02996) |
| 허리높이/어깨+ 엉덩이) | 0.481 (0.00092) | 0.48 (0.00219) | 0.48 (0.00236) | 0.476 (0.00504) | 0.475 (0.00583) |
| 허리높이/(어깨+ 엉덩이+무릎) | 0.402 (0.00071) | 0.401 (0.0018) | 0.4 (0.00151) | 0.398 (0.00343) | 0.397 (0.00482) |

표 3. 여성의 연령별 신체치수와 각 신체치수 대비 허리높이

표준편차가 상대적으로 가장 낮은 (어깨높이 + 엉덩이높이) 대비 허리높이 평균값을 사용하여 허리 위치를 예측하기로 하였다.

허리높이를 구하는 과정은 다음과 같다.

1. 어깨와 엉덩이, 그리고 발가락의 keypoint의 좌표를 구한다.
2. (발가락 keypoint의 y좌표 - 어깨 keypoint의 y좌표)를 구하여 어깨높이를 구한다.
3. (발가락 keypoint의 y좌표 - 엉덩이 keypoint의 y좌표)를 구하여 엉덩이높이를 구한다.
4. 표1과 표2에서 구한 비율과 (어깨높이+엉덩이높이)를 곱하여 허리의 y좌표를 구한다.

3.4.3 가슴 위치 추측

가슴 위치는 한국인의 평균 가슴높이에 대한 데이터가 충분하지 않기 때문에, 어깨 좌표에서 어깨높이의 약 10% 정도의 길이를 아래로 이동하여 가슴의 위치를 추정하였다.

3.4.4 배경 제거 및 길이 측정

초기에는 Canny Edge Detection을 사용하여 신체와 배경의 경계부분을 찾으려고 하였으나, 착용하고 있는 옷의 무늬 등으로 인해 측정이 어려워 Rembg 배경 제거 라이브러리를 사용하여 배경을 제거하였다. 배경을 제거한 후, 측정된 신체 부위의 좌표를 사용하여 길이를 측정하였다. 팔을 붙인 상태에서는 팔의 너비까지 허리와 가슴너비로 측정되기 때문에 그림 2와 같이 팔을 벌린 사진으로 측정하였다. 그림2는 [6] 데이터셋을 활용하였다.

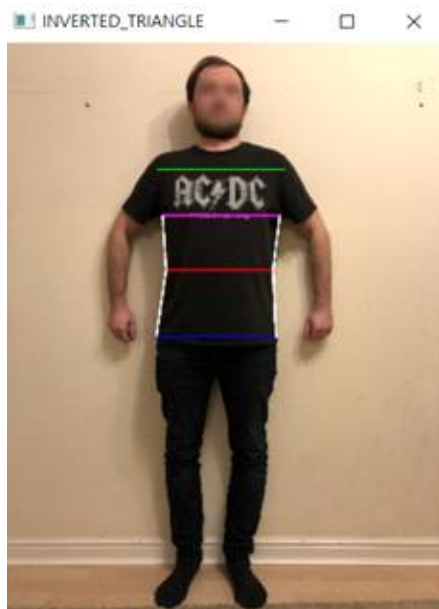


그림 2. 어깨, 가슴, 허리, 엉덩이 너비 측정

3.5 RandomForest를 활용한 추천시스템

3.5.1 데이터 전처리

데이터는 ai-hub에 연도별 패션 선호도 파악 및 추천 데이터의 2019년 데이터를 사용하였다. 해당 데이터는 옷 착용 이미지와 해당 이미지를 응답자가 평가한 json 형태의 자료로 구성되어있다. 평가 자료는 응답자 아이디, 평가한 이미지 이름, 패션 스타일, 이미지 성별, 평가내역, 응답자 정보로 구성되어있다. 각 사용자마다 선호하는 패션스타일은 아래 표6과 같이 패션스타일1~5로 정리되어있으며, 각 패션스타일마다 하나씩 골라 총 5개의 선호하는 패션스타일을 고르게 된다.

| 패션스타일 | 패션스타일1 | 패션스타일2 | 패션스타일3 | 패션스타일4 | 패션스타일5 |
|-------|-------------|--------------|----------|----------|----------|
| 1 | 화려하고 독특한 | 남성적 /여성적인 | 전통적인 | 포멀한 | 활발한 |
| 2 | 무난하고 평범한 | 중성적인 | 트렌디한 | 캐주얼한 | 점잖은 |
| 비고 | 1, 2중 택1 | 1, 2중 택1 | 1, 2중 택1 | 1, 2중 택1 | 1, 2중 택1 |

표 6. 응답자의 선호하는 패션스타일 선택

데이터 전처리 과정에서는 json 데이터를 csv로 변환하여 사용자 정보, 옷 정보, 평가 정보를 분리하였다. 옷에 대해 여러 사용자가 평가한 경우, 최빈값을 채택하는 방식으로 해당 옷의 최종 평가를 결정하였다. 이후 사용자가 선호하는 패션 스타일에 따라 퍼스널 컬러, 얼굴형, 체형을 유추하였다. 아래 표를 사용하여 스타일별로 퍼스널 컬러와 신체적 특성을 부여하고, 가장 빈도가 높은 특성을 최종 선택하였다.

| 선호하는 패션스타일 | 퍼스널컬러 | 얼굴형 | 체형 |
|------------|--------|--------------|-------------|
| 화려하고 독특한 | 봄, 겨울 | 타원형, 하트형 | 역삼각형, 사다리꼴형 |
| 무난하고 평범한 | 여름, 가을 | 둥근형, 사각형 | 둥근형, 직사각형 |
| 남성적인 | 가을, 겨울 | 긴형, 사각형 | 역삼각형, 사다리꼴형 |
| 중성적인 | 봄, 여름 | 타원형 | 직사각형 |
| 전통적인 | 가을, 겨울 | 둥근형, 사각형, 긴형 | 사다리꼴형, 둥근형 |
| 트렌디한 | 봄, 여름 | 하트형, 타원형 | 역삼각형, 삼각형 |
| 포멀한 | 가을, 겨울 | 긴형, 사각형 | 직사각형, 사다리꼴형 |
| 캐주얼한 | 봄, 여름 | 둥근형, 하트형 | 둥근형, 삼각형 |
| 활발한 | 봄, 여름 | 둥근형, 하트형 | 삼각형, 역삼각형 |
| 점잖은 | 가을, 겨울 | 사각형, 타원형, 긴형 | 사다리꼴형, 직사각형 |

표 7. 남성 응답자에 퍼스널 컬러, 얼굴형, 체형 부여

| 선호하는 패션스타일 | 퍼스널컬러 | 얼굴형 | 체형 |
|------------|--------|--------------|-------------|
| 화려하고 독특한 | 봄, 겨울 | 타원형, 하트형 | 역삼각형, 모래시계형 |
| 무난하고 평범한 | 여름, 가을 | 둥근형, 사각형 | 둥근형, 직사각형 |
| 남성적인 | 가을, 겨울 | 긴형, 사각형 | 모래시계형 |
| 중성적인 | 봄, 여름 | 타원형 | 직사각형 |
| 전통적인 | 가을, 겨울 | 둥근형, 사각형, 긴형 | 직사각형, 둥근형 |
| 트렌디한 | 봄, 여름 | 하트형, 타원형 | 역삼각형, 삼각형 |
| 포멀한 | 가을, 겨울 | 긴형, 사각형 | 직사각형, 모래시계형 |
| 캐주얼한 | 봄, 여름 | 둥근형, 하트형 | 둥근형, 삼각형 |
| 활발한 | 봄, 여름 | 둥근형, 하트형 | 삼각형, 역삼각형 |
| 점잖은 | 가을, 겨울 | 사각형, 타원형, 긴형 | 모래시계형, 직사각형 |

표 8. 여성 응답자에 퍼스널 컬러, 얼굴형, 체형 부여

3.5.2 Random Forest를 활용한 학습

학습에 사용된 데이터는 사용자 특성(퍼스널컬러, 얼굴형, 체형)과 옷의 특성을 feature로, 평점을 target으로 구성하였다. 데이터는 OneHot Encoding을 통해 전처리되었고, Scikit-Learn의 RandomForestRegressor를 사용하여 학습하였다. 학습 결과의 RMSE는 남성 데이터의 경우 훈련 데이터에서 **0.304**, 테스트 데이터에서 **0.759**로, 여성 데이터의 경우 훈련 데이터에서 **0.312**, 테스트 데이터에서 **0.773**으로 나타났다.

| 성별/데이터 | 훈련 데이터 | 테스트 데이터 |
|--------|--------|---------|
| 남 | 0.304 | 0.759 |
| 여 | 0.312 | 0.773 |

표 9. RandomForest를 사용한 훈련 결과

만들어진 모델을 사용하여 사용자 특성 - 옷 특성에 대해 평점 예측을 하여 평점이 높은 순서대로 추천을 해보면 표10의 결과가 나온다.

| 옷 이름 | 예측 평점(4점 만점) |
|---------------------------|--------------|
| W_81578_19_normcore_M.jpg | 3.780000 |
| W_63901_19_normcore_M.jpg | 3.770000 |
| W_54264_19_normcore_M.jpg | 3.760000 |
| T_17451_19_normcore_M.jpg | 3.760000 |
| W_81468_19_normcore_M.jpg | 3.750000 |
| W_01418_19_normcore_M.jpg | 3.743333 |
| T_19532_19_normcore_M.jpg | 3.740000 |
| T_17134_19_normcore_M.jpg | 3.700000 |
| W_31826_19_normcore_M.jpg | 3.700000 |
| T_17087_19_normcore_M.jpg | 3.683333 |

표 10. 20대 남성, 봄 퍼스널컬러, 둥근 얼굴형, 둥근 체형에 대한 평점 예측 결과

3.5.3 추천 품질에 대한 피드백

자문의견서에 기재된 추천 품질에 대한 피드백을 제공하는 것을 구현하였다. 구현방법은 3.5.2의 RandomForest를 사용한 모델을 사용하여 다음과 같이 구성하였다.

1. 추천 시 모델이 예측한 평점과 실제 사용자가 평가한 평점을 함께 반영하여 최종 평점을 산출한다. 모델의 예측 평점을 70%, 실제 사용자의 평점을 30% 반영하는 방식이다.
2. 사용자가 별점을 매기면 해당 데이터를 기존 훈련 데이터에 추가하고, 새로운 RandomForest 모델을 학습시킨다. 사용자가 매긴 평점은 사용자별로 기록되고, 실제 사용자의 평점에 반영된다.

표10의 추천결과를 3.5.3-1을 반영하여 추천하면 표11이 된다.

| 옷 이름 | 예상평점 | 사용자 평점 | 종합평점 |
|---------------------------|--------|--------|-------|
| W_81578_19_normcore_M.jpg | 3.78 | 4 | 3.846 |
| W_63901_19_normcore_M.jpg | 3.77 | 4 | 3.839 |
| T_17451_19_normcore_M.jpg | 3.76 | 4 | 3.832 |
| W_54264_19_normcore_M.jpg | 3.76 | 4 | 3.832 |
| W_81468_19_normcore_M.jpg | 3.75 | 4 | 3.825 |
| T_19532_19_normcore_M.jpg | 3.74 | 4 | 3.818 |
| W_31826_19_normcore_M.jpg | 3.7 | 4 | 3.79 |
| T_17134_19_normcore_M.jpg | 3.7 | 4 | 3.79 |
| T_17087_19_normcore_M.jpg | 3.6833 | 4 | 3.778 |
| T_19542_19_normcore_M.jpg | 3.68 | 4 | 3.776 |

표 11. 20대 남성, 봄 퍼스널컬러, 둥근 얼굴형, 둥근 체형에서 평점 예측

이제 표11에서 추천시스템이 가장 추천하는 옷에 대하여 평점을 1점으로 평가한 후, 다시 추천하면 표12와 같다.

| 옷 이름 | 예상평점 | 사용자 평점 | 종합평점 |
|---------------------------|-------|--------|--------|
| W_54264_19_normcore_M.jpg | 3.77 | 4 | 3.839 |
| T_17451_19_normcore_M.jpg | 3.77 | 4 | 3.839 |
| W_63901_19_normcore_M.jpg | 3.76 | 4 | 3.832 |
| W_81468_19_normcore_M.jpg | 3.76 | 4 | 3.832 |
| W_31826_19_normcore_M.jpg | 3.705 | 4 | 3.7935 |
| T_17134_19_normcore_M.jpg | 3.7 | 4 | 3.79 |
| W_17682_19_normcore_M.jpg | 3.68 | 4 | 3.776 |
| W_52075_19_normcore_M.jpg | 3.67 | 4 | 3.769 |
| W_25216_19_normcore_M.jpg | 3.67 | 4 | 3.769 |
| W_92952_19_normcore_M.jpg | 3.67 | 4 | 3.769 |

표 12. 20대 남성, 봄 퍼스널컬러, 둥근 얼굴형, 둥근 체형에서 피드백 후 추천 결과

3.6 콘텐츠 기반 협업필터링 추천시스템

3.6.1 데이터 전처리

데이터는 AI-hub의 K-Fashion 이미지를 사용하였다. 해당 데이터는 120만 장의 옷 착용 사진과 해당 옷의 특징을 json 파일로 정리되어있다. 이를 파싱하여 상의, 하의, 아우터, 원피스 등의 카테고리로 구분하고, csv 파일로 정리하였다.

3.6.2 Tf-Idf Vectorizing

Tf-Idf Vectorizing은 키워드의 중요도를 나타내는 기법으로, 옷의 각 특성을 하나의 문자열로 합친 후 Scikit-Learn의 Tf-Idf Vectorizer를 사용하여 벡터화하였다. 이 과정에서 L2 정규화를 적용하여 과적합을 방지하였다.

3.6.3 코사인 유사도 계산 및 옷 특성 연결

코사인 유사도는 내적공간의 두 벡터간 각도의 코사인 값을 이용하여 측정된 유사도이다. $X=(x_1,...,x_n)$, $Y=(y_1,...,y_n)$ 을 두 문서의 벡터쌍이라고 하면 코사인 유사도는 아래와 같이 구할 수 있다.

$$Cosine(X, Y) = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2} \sqrt{\sum y_i^2}}$$

Scikit-Learn의 cosine_similarity 함수를 사용하여 벡터화된 옷 특성 간의 유사도를 계산하고, 유사도가 높은 순서대로 추천하였다. 이후, 사용자의 퍼스널컬러, 얼굴형, 체형에 어울리는 옷 특성을 연결하였다. 예를 들어 사각형 얼굴에는 어울리는 상의 특성에 라운드넥을 추가하였다.

| 옷번호 | 스타일 | 옷특징 | 유사도 |
|---------|------|---|--------|
| 1038491 | 스트리트 | 카키 긴팔 스트리트 스포티 티셔츠 저지 그래픽 오버사이즈 롱 라운드넥 | 1 |
| 37620 | 스트리트 | 카키 긴팔 화이트 스트리트 티셔츠 저지 그래픽 오버사이즈 롱 라운드넥 | 0.8614 |
| 1036139 | 스트리트 | 카키 긴팔 스트리트 스포티 티셔츠 우븐 그래픽 루즈 롱 라운드넥 | 0.8309 |
| 190270 | 스트리트 | 스트리트 스포티 티셔츠 저지 그래픽 오버사이즈 롱 | 0.8276 |
| 1213067 | 스트리트 | 카키 긴팔 스트리트 스포티 티셔츠 드롭숄더 저지 무지 오버사이즈 | 0.8208 |
| 735440 | 스트리트 | 블랙 긴팔 화이트 스트리트 스포티 티셔츠 저지 그래픽 오버사이즈 롱 라운드넥 | 0.8035 |
| 735430 | 스트리트 | 블랙 긴팔 화이트 스트리트 스포티 티셔츠 저지 그래픽 오버사이즈 롱 라운드넥 | 0.8035 |
| 735410 | 스트리트 | 블랙 긴팔 화이트 스트리트 스포티 티셔츠 저지 그래픽 오버사이즈 롱 라운드넥 | 0.8035 |
| 37630 | 스트리트 | 카키 긴팔 화이트 스트리트 티셔츠 슬릿 저지 그래픽 오버사이즈 롱 라운드넥 | 0.7719 |
| 397965 | 스트리트 | 카키 긴팔 블랙 스트리트 티셔츠 드롭숄더 저지 레터링, 그래픽 오버사이즈 롱 라운드넥 | 0.7459 |

표 13. '카키 긴팔 스트리트 스포티 티셔츠 저지 그래픽 오버사이즈 롱 라운드넥' 특성과 유사한 상의

4. 연구 결과 분석 및 평가

4.1 퍼스널 컬러 측정 모델

퍼스널 컬러 측정은 사용자의 피부색을 기반으로 K-means 클러스터링을 활용하여 네 가지 계절(봄, 여름, 가을, 겨울)로 분류되었다. 각 계절에 따라 피부색의 HSV(Value, Saturation) 및 Lab(b 채널) 값을 사용하여 퍼스널 컬러를 분류하였으며, 그림에 나타난 바와 같이, 각 클러스터에서 특징적인 색상 값을 도출하였다.

아래 그림3은 각 계절(클러스터)의 주요 색상 특징을 시각화한 그래프이다. 이 그래프는 Value(명도), Saturation(채도), b 채널(황-청색축) 값의 평균을 나타낸다.

그래프에서 알 수 있듯, 봄 타입은 Value와 Saturation 값이 높아 명도와 채도가 높은 따뜻한 색조를 특징으로 한다. 여름 타입은 상대적으로 Saturation 값이 낮아, 쿨톤에 해당하는 색상들이 주로 나타났다. 가을 타입은 b 채널 값이 높아 따뜻한 색조가 특징이며, 겨울 타입은 Value가 비교적 낮고 Saturation 값이 중간 정도로 나타나 명도는 낮지만 고채도의 차가운 색상을 반영하고 있다.

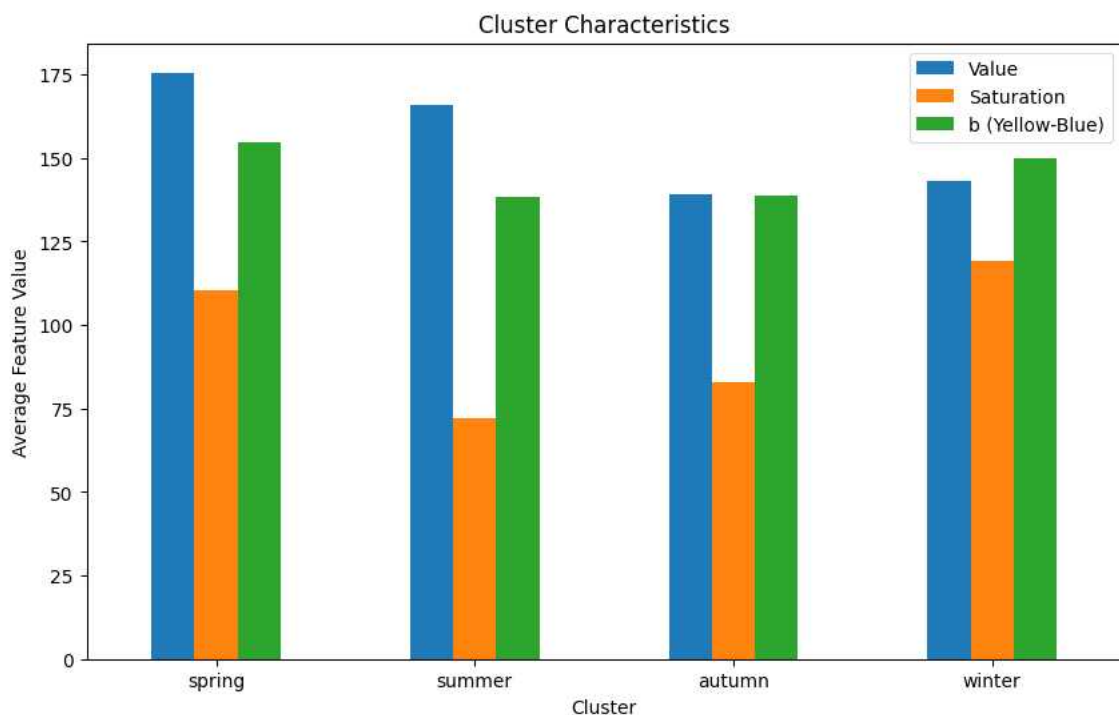


그림 3. 계절별 클러스터

4.2 얼굴형 측정 모델

4.2.1 EfficientNetB4를 제외한 나머지만 학습

사전학습된 EfficientNetB4의 가중치는 수정하지 않고, 나머지 부분만 학습하도록 하였다. 25 epoch 학습시켰으며, adam optimizer(learning rate = 0.001)를 사용하였다. 마지막 에포크에서 훈련 정확도는 **0.9851**, 테스트 정확도는 **0.6961**이 나왔다.

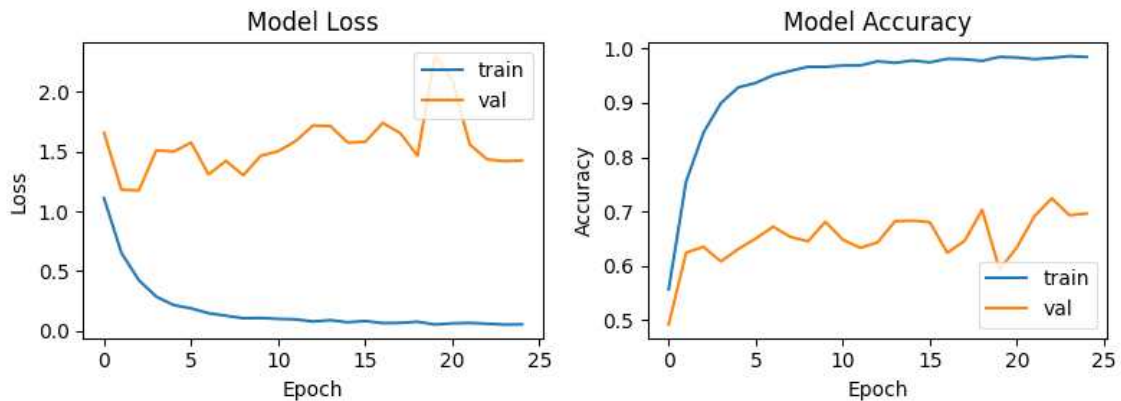


그림 4. EfficientNetB4를 제외한 나머지 학습 그래프

4.2.2 전체 학습

EfficientNetB4를 제외하고 학습한 후에, EfficientNetb4를 포함하여 학습하였다. 기존 학습된 가중치의 과도한 수정을 방지하기 위해 학습률은 0.0001로 설정하였다. 마지막 에포크에서, 훈련 정확도는 **0.9993**, 테스트 정확도는 **0.7513**가 나왔다. EarlyStopping을 적용하였고, restore_best_weights=True로 설정하였기 때문에 훈련 정확도 **0.9994**, 테스트 정확도 **0.7583** 모델이 저장되었다.

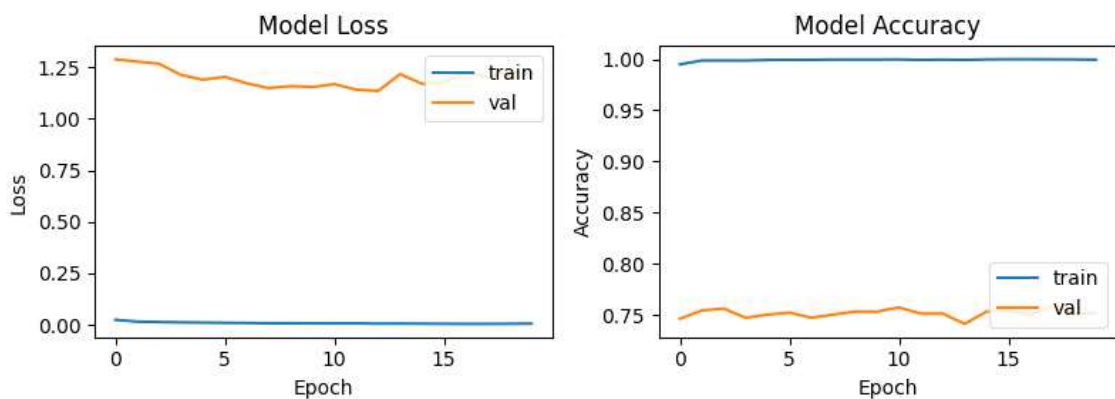


그림 5. EfficientNetB4를 포함한 학습 그래프

4.3 체형 측정 모델

측정된 어깨 너비, 가슴너비, 허리너비, 엉덩이너비에 따라 다음과 같이 분류하였다.

| 기준 | 분류 |
|--|-------|
| 어깨너비>엉덩이너비 and 엉덩이너비>허리너비 and 가슴너비>엉덩이너비 | 사다리꼴형 |
| 허리너비>어깨너비 and 허리너비>엉덩이너비 | 동근형 |
| 어깨너비>엉덩이너비 and 어깨너비>가슴너비 | 역삼각형 |
| 엉덩이너비>어깨너비 and 엉덩이너비>가슴너비 | 삼각형 |
| 그외 | 사각형 |

표 4. 남성의 체형 분류 기준

| 기준 | 분류 |
|---------------------------|-------|
| 어깨너비>허리너비 and 엉덩이너비>허리너비 | 모래시계형 |
| 허리너비>어깨너비 and 허리너비>엉덩이너비 | 동근형 |
| 어깨너비>엉덩이너비 and 어깨너비>가슴너비 | 역삼각형 |
| 엉덩이너비>어깨너비 and 엉덩이너비>가슴너비 | 삼각형 |
| 그외 | 사각형 |

표 5. 여성의 체형 분류 기준

4.4 추천시스템

RandomForest 기반 학습 결과, 남성 데이터의 훈련 RMSE는 **0.304**, 테스트 RMSE는 **0.759**로 나타났으며, 여성 데이터의 경우 훈련 RMSE는 **0.312**, 테스트 RMSE는 **0.773**으로 기록되었다. 이는 추천 모델이 사용자의 신체적 특성과 옷의 특성을 정확히 반영하여 패션 아이템을 추천할 수 있음을 의미한다. 또한, 콘텐츠 기반 협업 필터링 추천시스템은 옷의 텍스트 특성을 바탕으로 비슷한 옷을 추천하는 데 효과적이었으며, 이를 통해 사용자에게 어울리는 스타일의 다양한 옷을 탐색할 수 있었다.

패션 추천 시스템

얼굴 사진:
[파일 선택] 선택된 파일 없음

전신 사진:
[파일 선택] 선택된 파일 없음

팔을 올린 전신 사진:
[파일 선택] 선택된 파일 없음

성별:
성별을 선택하세요

나이:

계절 (선택 사항):
선택하지 않음

추천 받기 신체 정보 직접 입력하기

패션 추천 시스템

패스워드 컬러:
선택하세요

얼굴형:
선택하세요

체형:
성별을 먼저 선택하세요

성별:
성별을 선택하세요

나이:

계절 (선택 사항):
선택하지 않음

추천 받기 사진으로 분석하기

그림 6. 패션 추천을 위한 정보 입력 페이지

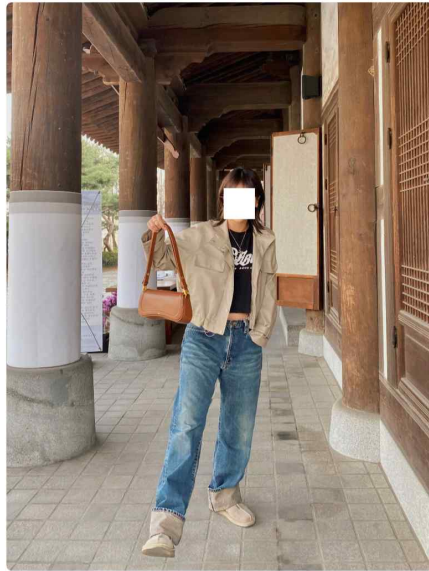
신체적 특징 분석 결과

퍼스널 컬러: summer

얼굴형: oval

체형: round

추천 패션 아이템



예상 평점: 3.63
평균 평점: 4.00
총 점수: 3.74

스타일: normcore

어울리는계절: 봄/가을

적합한착용상황: 데이트

핏평가: 적당함

색감: 밝음

색감: 따듯함

분위기: 가벼움

특징:

트렌디하다, 깔끔하다, 실용적이다, 활동적이다, 편안하다, 여성적이다



예상 평점: 3.62
평균 평점: 4.00
총 점수: 3.73

스타일: normcore

어울리는계절: 봄/가을

적합한착용상황: 출근

핏평가: 헐렁함

색감: 밝음

색감: 따듯함

분위기: 가벼움

특징:

트렌디하다, 깔끔하다, 무난하다, 실용적이다, 활동적이다, 편안하다, 발랄하다, 여성적이다

그림 7. 신체적 특징 분석 결과 및 패션 추천 페이지

5. 결론 및 향후 연구 방향

퍼스널 컬러, 얼굴형, 체형이라는 사용자 신체적 특성을 기반으로 맞춤형 패션을 추천하는 시스템을 구축하여 개인화 서비스의 가능성을 확인하였다. 특히, RandomForest 기반 추천 시스템은 높은 예측 정확도를 기록했으며, 사용자 피드백을 반영한 추천 품질 개선 기능이 시스템의 성능을 지속적으로 높이는 데 중요한 역할을 했다.

이러한 결과는 사용자 맞춤형 서비스의 중요성과 개인화 추천 시스템이 패션 산업에서 어떻게 적용될 수 있는지에 대한 가능성을 보여주었다.

향후 다양한 데이터를 확보하고, 실시간 피드백을 반영한 추천 시스템의 고도화를 통해 더 나은 사용자 경험을 제공할 수 있을 것이다. 또한, 딥러닝 기반의 고도화된 추천 알고리즘과 패션 트렌드 분석을 통해 사용자에게 더욱 개인화되고 실시간으로 반응하는 패션 추천 시스템을 구축한다면 패션 산업이 더욱 디지털화되고 사용자 맞춤형 서비스에 대한 수요에 충족할 것으로 기대된다.

추가적으로, 쇼핑몰과의 연동을 통해 추천된 패션 아이템을 직접 구매할 수 있는 연계 시스템을 구축함으로써 사용자 편의성을 증대시키고, 패션 추천이 단순히 스타일 제안을 넘어 실제 의류 구매로 이어지는 통합된 사용자 경험을 제공할 수 있을 것으로 기대된다.

6. 참고 문헌

- [1] Yun-Seok Jung, "A Study on the Quantitative Diagnosis Model of Personal Color," *Journal of Convergence for Information Technology*, Vol. 11, No. 11, pp. 277-287, 2021. (in Korean)
- [2] Jong-Suk An, "A Study on Effective Image Making Depending on Hair Style and Neckline," *J Korean Soc Cosmetol*, Vol. 15, No. 1, pp.342-351, 2009. (in Korean)
- [3] Soo-ae Kwon, *Fashion and Life*, Gyohakyungusa, 2016.
- [4] "2023 Consumption Trend Series - 03 Personalized Services," MezzoMedia [Online],
Available: https://www.mezzomedia.co.kr/data/insight_m_file/insight_m_file_1605.pdf
(downloaded 2024, May. 19)
- [5] So-young Lee, "Personal Color Tone Type and Categorization of Harmonious Colors According to Skin Color," M.S. thesis, Graduate School of Cultural and Information Policy, Hongik Univ., Seoul, South Korea, 2019. (in Korean)
- [6] Cameron Patrick Trotter et al. Human Body Shape Classification Dataset. https://data.ncl.ac.uk/articles/dataset/Human_Body_Shape_Classification_Dataset/19307300?file=34292915